DOCUMENT RESUME

ED 478 018 FL 027 750

AUTHOR	Kirkwold, Lorne O.
TITLE	Issues in Course Development, Evaluation, and Testing: A Case Study from Japan.
PUB DATE	2003-06-04
NOTE	10p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Communication Skills; *Curriculum Development; *English (Second Language); Evaluation Methods; Foreign Countries; Grammar ; Higher Education; *Language Proficiency; Second Language Instruction; Teaching Methods; *Test Construction
IDENTIFIERS	Conversation; Japan; Universal Grammar

ABSTRACT

This study evaluated one professor's teaching of English as a Second Language courses at three Japanese universities. Instruction was modified to make the courses similar communicatively to one another. The professor administered a test, found in a journal, at the three universities, comparing results to evaluate the appropriateness of his own teaching within each class. The instrument was originally developed to investigate the nature of pro-drop transfer of Spanish speakers learning English at McGill University in Montreal, Quebec, Canada. It consisted of 31 items, each requiring a grammaticality judgment. Results indicated that giving a similar test to students at three different universities was a useful way of assessing teaching and confirming that the teaching was on target. The instrument was valid and showed promise as a measure of proficiency. (Contains 13 references.) (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 478 018

Issues in Course Development, Evaluation, and Testing:

A Case Study from Japan

Lorne O. Kirkwold

Abstract

This article describes an evaluation I conducted of courses I was teaching part-time by comparing the results of a test found in a journal which I gave at three prominent universities in the Sapporo region. I changed the instruction of the courses I evaluated to make them more similar communicatively to the content of the courses I was giving at the other universities. Qualitatively, the instrument, although developed to investigate L_1 pro-drop parameter influence by the original author, shows promise for validation as a measure of proficiency where other testing may produce less exacting scores. The article is intended for readers interested in communicative course development, testing, and course evaluation at Japanese universities.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Lorne O. Kirkwold

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

FL027250

Issues in Course Development, Evaluation, and Testing: A Case Study from Japan

Lorne O. Kirkwold

Hokkai-Gakuen University, Sapporo, Japan

This article describes an evaluation I conducted of courses I was teaching part-time by comparing the results of a test found in a journal which I gave at three prominent universities in the Sapporo region. I changed the instruction of the courses I evaluated to make them more similar communicatively to the content of the courses I was giving at the other universities. Qualitatively, the instrument, although developed to investigate L_1 pro-drop parameter influence by the original author, shows promise for validation as a measure of proficiency where other testing may produce less exacting scores. The article is intended for readers interested in communicative course development, testing, and course evaluation at Japanese universities.

Introduction

There are various approaches to course development. A traditional means of elaborating students' needs is to frame objectives linguistically. This can involve the comparison of linguistic structures in the two languages and evaluating the difficulty of the structures to be practiced in the foreign language. Those structures that are most different between the two languages will presumably require the most attention. Lado's (1957) work is one of the best known discussing such an analysis for developing objectives. In more recent curriculum innovations, objectives are framed around communicative tasks, sometimes these tasks are the results of a needs analysis. Nunan (1989) describes communicative tasks aimed at developing fluency. Long (1985) considers how such tasks can be graded.

Course development case study

Often when we have decisions about program development, we have a chance to exchange our ideas with our colleagues. Indeed, this has been my experience at Hokkai-Gakuen University (HGU), my full-time job, where we have developed a regular stream of skill-based courses in Reading, Writing, Speaking, and Listening for students to pursue during their first two years of studies in our department of English language and culture. In these same skill areas are parallel courses in the seminar stream, ideally for students who demonstrate greater aptitude. This brings me to the problem which I would like to describe as the focus for this paper. In addition to my regular job at HGU,

a private institution in Sapporo, I also have two part-time jobs outside: the first at Hokkaido University (HU), the city's national university; the second at Otaru University of Commerce (OUC), also a national university within commuting distance. In Japan, a foreign-language course at a university is given during a ninety-minute period with meetings scheduled once a week. Depending on the institution, there will be 12 to 15 classes per semester. It is usual for classes to meet a full year for a total of 24 to 30 classes, but some language courses run only a semester. At my regular job, I have worked closely with my colleagues in determining what the objectives should be in the courses we teach. These are reflected in the descriptions we have written jointly for the language-skill courses of a year's length and the department's list of textbooks which we have chosen together. At my part-time job at HU, two of the colleagues in my usual department are employed at the same time. We collaborate together to determine what we are going to teach in English *(Eigo)* I. This is in addition to objectives and requirements set in English by HU faculty members. It is a general one-semester course, whose broad objective is conversational speaking, in which we have had enrolled day students from education, law, medicine, dentistry, engineering, agriculture, and veterinary. Going to my second part-time job, although I may see other colleagues and "talk shop," I have not found the same amount of formal exchange about the objectives for English courses at OUC. My assignment there is English I in the daytime and English II for the night students. These courses run the full year.

I began my paper by mentioning the alternatives for the development of a language program of either linguistic or communicative objectives. As for my selection on my own at OUC, I chose the textbook *Clear Speech*. Readers interested in the course design are referred to the *Teacher's Resource Book* (Gilbert, 1993). It is a compromise I see between these poles, and I use this book for both courses there. The text written at an earlier time may have been critiqued as an audio-lingual method by Rivers (1964). Certainly there is opportunity for repetition using the accompanying recording. Dialogues are also modeled tightly. In the early chapters, dictation exercises are the focus of listening practice. Gilbert does not prescribe memorization though. The content of the course is a systematic presentation of the problematic sounds of the language (for example, voiced and unvoiced **th**, **b/v**, the plural **-s**, and **-ed**) in addition to a revision of stress, rhythm, linking and contractions. These features are generally explained and modeled by the teacher or the program on the tape. After this, there

may be some choral repetitions. Then students practice the features they have revised in pair work, essentially short questions and answers. The material seems to assume some previous study of the spoken features of the language, and in this regard, the textbook accommodates a large number of students readily. I suspect they have studied English in similar ways in high school.

This is the position from which I started, but I was more satisfied with the greater focus on the communicative aspect of language learning both at HU and HGU. For example, in speaking courses at these universities, we use *Marathon Mouth* and *Marathon Mouth Plus*. Readers interested in the course design are referred respectively to *Marathon MouthTeacher's Edition* (Koustaff, Gaston, & Shimizu, 2000) and *Marathon Mouth Plus Teacher's Edition* (Shimizu & Gaston, 2000). The pair work in a chapter of this textbook series typically consists of information gap activities thematically developed. Afterwards, the focus of the recordings is often listening for specific information. The interview activity which comes near the end of the chapter allows students to exchange information of personal interest among themselves and with the teacher. These exchanges can be cultural. In Speaking Seminar 2 at HGU the year I gathered data for this study, I used *Amazing! Interviews and Conversations* (Bates, 1993). The recorded material consists of interviews, albeit scripted, developed around human interest stories in Canadian newspapers. Not only are there comprehension exercises focusing on specific details and note-taking activities, the material also offers ample opportunity for instruction about Canadian culture. In contrast, teaching my lessons at OUC, I felt the instruction was "hot and cold." I was not certain either as to the level of students relative to those at the other two institutions, where I had been teaching for a longer period of time. I decided at the beginning of the school year in 2001 that I would administer the same test to all classes at the three institutions in order to compare levels of proficiency.

The instrument and results

The instrument I chose was a test developed for other purposes in the universal grammar (UG) literature. What aroused my curiosity was testing the pro-drop parameter in a Japanese context. White (1985) used the instrument for speculating about the nature of pro-drop transfer of Spanish speakers learning English at McGill

University in Montreal, Canada.[1] The test consists of thirty-one items, each requiring a grammaticality judgment. Six test missing subjects. There are 13 Wh- questions. Interspersed among the questions are four sentences with affirmative intent whose subjects and verbs are ungrammatically inversed (presumably following allowable Spanish order) and a fifth with stylistic inversion following an existential "there." Six items, of which five are examples of relative clauses, test subordination. Finally, there is one example of the SVC structure. A student takes the test by indicating grammaticality in the following way. Marking OK for a given item indicates the student believes the sentence is correct. In the case of incorrect sentences, the student is intended to correct the mistake(s) and copy the corrected version on the line below. In order to score a point, a student must indicate OK appropriately or, in the case of an incorrect sentence, a suitable correction must be provided. There are no partial points. Although there are 31 items on the original test, when I scored the test myself, I excluded 6 items. A perfect score would therefore be 25 in my results. This is the greatest modification I made in the administration of the test compared to White's description. I administered the test to: the English I classes at HU on April 11, 2001 and October 3, 2001; and the English I and II classes at OUC on October 4, 2001. HGU students were tested between April 12 and 17, 2001, with the exception of Speaking Seminar 2 students, who were tested November 14, 2001. Students were in attendance with only a few exceptions; generally the absentees took the test at other times. The results follow below in rank order:

Score	U	Course	n	Yr	D/N	Major	Text
16.69	HU	English I (64)	32	1	D	Agriculture & Veterinary	*MM+*
16.56	HGU	Writing Seminar 3	9	3	D	English	
16.10	HU	English I (35)	30	1	D	Medical / Dental	*MM*
14.93	HU	English I (43)	29	1	D	Engineering	*MM+*
14.69	HGU	Speaking Seminar 2	13	3	N	English	*Amazing!*
14.22	OUC	English I (E-133B)	32	1	D	Commerce	*CS*
14.17	HU	English I (9)	29	1	D	Education / Law	*MM*
13.38	OUC	English II (E-23C)	45	2	N	Commerce	*CS*
12.88	HGU	Speaking 2	24	2	D	English	*MM+*
12.86	HGU	Speaking/Listening Sem	14	1	D	Japanese	*MM*

[1] I extend my thanks to Professor White for allowing me to use her instrument.

12.29	HGU	Writing Seminar I	17	2	N	English
12.10	HGU	Reading 1	29	1	D	English
11.68	HGU	Listening 1	22	1	D	English
11.59	HGU	Reading 1	22	1	N	English

Beside each score is given the university, course (with section number indicated in parentheses), and number of students, along with information about the students' year, time at which they study (D for 'day' and N for 'night'), and major. Lines containing OUC results are highlighted in bold font. For the purpose of making comparisons among the speaking courses, I have indicated textbooks: *MM* for *Marathon Mouth, MM+* for *Marathon Mouth Plus, C S* for *Clear Speech, Amazing!* for *Amazing! Interviews and conversations.*

Discussion

Describing my ranking qualitatively is the first step towards the validation of the test for this purpose. Students at Hokkaido University occupy the upper ranks. There are two third-year Hokkai-Gakuen seminars for English majors that rank among the Hokkaido University results. The OUC day students rank one place above the lowest results for Hokkaido University; the OUC night students, one place below. The second-year English majors at HGU follow. There are a day class and a night class, but with the first-year Japanese majors' results falling between. In other words, second-year daytime HGU English majors placed higher than did first-year daytime Japanese Studies students, who placed higher than second-year nighttime English majors. The last three results are those of first-year HGU English majors, with the night students ranking the lowest.

In brief, my interpretation of these results is that the two courses at OUC are both in the middle third, in terms of proficiency, of my teaching assignments. I can use these data to make a case for the desirability of switching the focus in these classes. We continued to use the textbook, but for approximately 30 minutes each class. In the remaining hour, I pursued activities which I believed to challenge the students more in terms of their "communicative competence" (Canale & Swain, 1980). Such activities included watching videos developed for EFL instruction; for example, *Laugh and Learn with Mr. Bean* (Hamada & Akimoto, 2001), *Lost Secret 2000* (O'Neill, 1996) and *Family Album, USA* (Kelty, 1991). The authors of these materials have their own

design; however, I used these videos for their culture and storyline in the following way. We would view an entire episode at one time. I would then direct a structured production task. Students were to volunteer one or two sentences about what they had seen. I would write these sentences on the board, and then students would do a writing assignment based on the viewing of one (or more) of these episodes. My satisfaction with the change was such that I taught the entire 2002 - 2003 year in this way. At the end of that course, three students wrote these comments in English in an evaluation whose results were returned to me by OUC:

"Video assignment is very nice for me. I wanna be more good listener of English, so this class is good for me. It's important for me to listen and speak English. This is much for me."

"I enjoyed this class. To watch many videoes is good."

"This class is very interesting. The teacher used video much. The teacher speaks only English."

Evaluation of my study and directions for further research

An exercise such as this one may be useful for others who are in a similar situation. Many foreign-language teachers in Japan teach part-time but likely have little time to collaborate with colleagues. I have found giving a similar test to students at three different universities a helpful means of assessing an appropriate level in my teaching or confirming that I am on target. The comparison I have presented has allowed me to make decisions with more confidence about the nature of the material I can present to a certain group of students. What I believe to be novel about my research is that I have found in the literature an instrument suitable for analysis while conducting my own evaluation. Researchers develop instruments but it is rare to see discussion of their suitability by others as articles in language learning journals. Discussion of validity may well be abundant, but it is usually the same researcher who ascertains the suitability of his or her instrument. I believe disseminating a developing instrument through the journals will enhance its validity by increasing reliability. Colleagues in various places can confirm that the instrument is providing useful assessments by presenting data such as those I have gathered. To this end, a relatively objective test, like the one I have used, may ensure a certain level of scoring consistency. The bands of communicative tests may include descriptors which are not apt to be interpreted in the same way by native-

speaking and non-native speaking colleagues. Looking for a suitable means for assessment, we may consider the need for a "finer" instrument required in situations like my own where the relative homogeneity of the students would otherwise result in placing the majority in the same proficiency band.

At the same time, there is a body of literature associated with instruments described in journals. This may provide insight to the classroom teacher who is trying to understand how students may be learning and why particular points are problematic. In UG literature, the description of parameters may assist the teacher in understanding the development of non-mother tongue learning.

Caveats based on this experience

A journal article will not completely describe all steps required in the development of the instrument. In order to assist our colleagues, we should document with explanation amendments we make. For example, I had to add directions and examples to the instrument before I administered it. I also tried out the test myself in order to determine how I would score the results. Then I discovered that some familiarity with the students' responses before scoring is desirable in determining allowable correct answers. Consider item #5: "The policeman didn't know when did escape the prisoner." The expected correction is "The policemen didn't know when the prisoner escaped," and "The policemen didn't know when the prisoner did escape" is acceptable too. On the other hand, the students' correction, "The policemen didn't know when the prisoner had escaped" requires some judgment for the scoring. I have decided not to allow it because it is not the simplest correction. Although I administered the 31 items as they appear in the reference, I came to exclude from scoring the six items as numbered in the appendix (White, 1985, p. 62) which follow: #4, 11, 20, 23, 24, and 29. As for #4, White mentions herself that the pronoun "it" in this item could be omitted in spoken English. This makes scoring the item ambiguous. Similarly, #11, with its existential "there," seems plausible for students familiar with poetry or literature. White, however, intended this item as ungrammatical. Similar discussion could be written about the other items I have excluded from scoring. In these instances, native-speaker comparisons would be useful. The breakdown may prove interesting: some would be acceptable to a certain number of native speakers. In such cases, the explanation may be related to variation in dialects. In this train of thought, answer keys annotated with allowable

answers and discussion even in the article itself about how the wrong answers came to be excluded may facilitate the development of a testing instrument in various places. Explanation shared in this way may be of interest to readers wishing to experiment with the instrument.

References

Bates, S. (1993). *Amazing! Interviews and conversations. Instructor's manual.* Scarborough, ON: Prentice-Hall Canada Inc. (ISBN: 0-13-014390-1)

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1,* 1 - 47.

Gilbert, J. B. (1993). *Clear speech: Pronunciation and listening comprehension in North American English. Teacher's resource book* (2nd ed.). Cambridge: Cambridge University Press. (ISBN: 0-521-42116-0)

Hamada, M., & Akimoto, H. (2001). *Laugh and learn with Mr. Bean.* Tokyo: MacMillan LanguageHouse Ltd. (ISBN: 4-89585-390-X C3082)

Kelty, J. (1991). *Family album USA (Book 1): Teacher's edition.* New York: Maxwell Macmillan International Publishing Group. (ISBN: 0-02-332810-X)

Koustaff, L., Gaston, B., & Shimizu, P. (2000). *Marathon mouth teacher's edition* (5th ed.). Fukuoka, Japan: Intercom Press, Inc. (ISBN: 4-900689-11-4 C0082)

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor, MI: The University of Michigan Press.

Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77 - 99). Clevedon, UK: Multilingual Matters Ltd.

Nunan, D. (1989). *Designing tasks for the communicative classroom.* Cambridge: Cambridge University Press.

O'Neill, R. (1996). *Lost secret 2000: Teacher's guide.* Tokyo: Meynard Publishing Ltd. (ISBN: 4-944025-50-5 C1082)

Rivers, W. (1964). *The psychologist and the foreign-language teacher.* Chicago: The University of Chicago Press.

Shimizu, P., & Gaston, B. (2000). *Marathon mouth plus teacher's edition* (2nd ed.). Fukuoka: Intercom Press, Inc. (ISBN: 4-900689-25-4)

White, L. (1985). The "pro-drop" parameter in adult second language acquisition. *Language Learning, 35,* 47-62.

Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Issues in Course Development, Evaluation, and Testing: A Case Study from Japan	
Author(s): Lorne O. Kirkwold	
Corporate Source:	Publication Date: 2003 - 6 - 4

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

☐ *

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Lorne O. Kirkwold	Printed Name/Position/Title: Lorne O. Kirkwold, EFL Lecturer	
Organization/Address: Hokkai-Gakuen University, 4-1-40 Asahi-machi, Toyohira-ku, Sapporo 062-8605 JAPAN	Telephone: +81-11-841-1161	FAX: +81-11-824-7729
	E-Mail Address: kirkwold@jin.hokkai-s-u.ac.jp	Date: 2003-6-4

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on
Languages & Linguistics
4646 40TH ST. NW
WASHINGTON, D.C. 20016-1859

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 9/97)
PREVIOUS VERSIONS OF THIS FORM ARE OBSOLETE.